



## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **09081175 A**(43) Date of publication of application: **28.03.97**

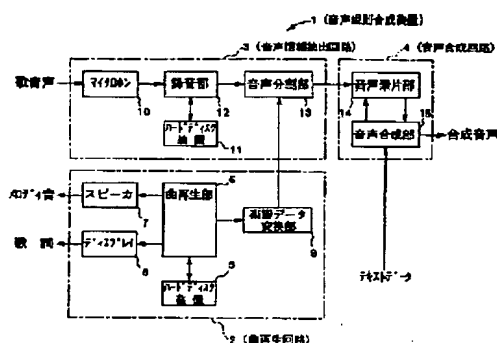
(51) Int. Cl. **G10L 3/00**  
**G10L 5/04**

(21) Application number: **07262473**(71) Applicant: **TOYO COMMUN EQUIP CO LTD**(22) Date of filing: **14.09.95**(72) Inventor: **FURUKAWA SUKEYUKI****(54) VOICE RULE SYNTHESIS DEVICE****(57) Abstract:**

**PROBLEM TO BE SOLVED:** To facilitate the generation of voice piece and to easily change the speaker of synthesized voice without requiring a large scale computer by conducting a necessary labeling process while generating voice pieces and constituting a circuit which process voice signals and generates voice information and a circuit which performs voice synthesis on one computer system.

**SOLUTION:** The lyric lines of a song which is beforehand registered by a music reproducing circuit 2 are displayed and melody sound is outputted while making a user sing the song. Then, a voice information extracting circuit 3 fetches the singing voice of the use, divides the singing voice for every sound syllable based on the information such as the length of each note of the singing voice and the song as well as the tempo and puts labels for every sound syllable to generate voice pieces. When text data, which are the object of voice synthesis, are inputted, the voice pieces are combined in accordance with the data and synthesized voice is generated.

COPYRIGHT: (C)1997,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-81175

(43) 公開日 平成9年(1997)3月28日

(51) Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L	3/00		G 1 0 L	H
	5/04			F

審査請求 未請求 請求項の数 2 F D (全 5 頁)

(21) 出願番号 特願平7-262473

(22) 出願日 平成7年(1995)9月14日

(71) 出願人 000003104

東洋通信機株式会社

神奈川県高座郡寒川町小谷2丁目1番1号

(72) 発明者 古川 祐行

神奈川県高座郡寒川町小谷2丁目1番1号

東洋通信機株式会社内

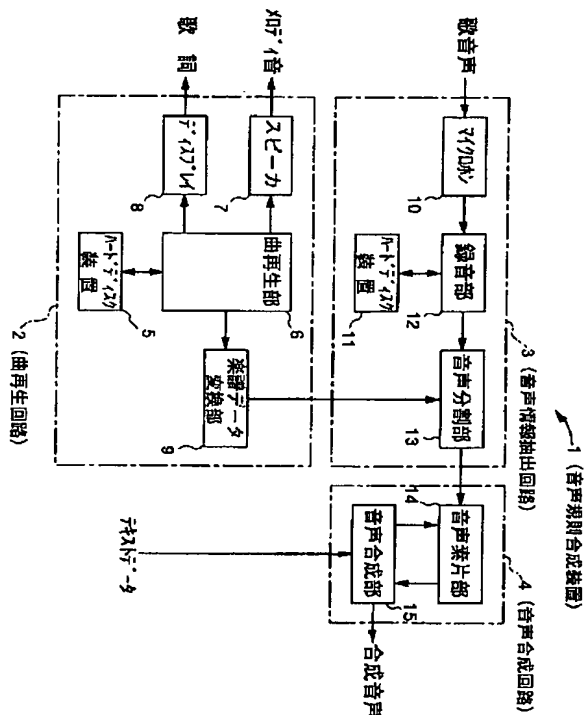
(74) 代理人 弁理士 鈴木 均

#### (54) 【発明の名称】 音声規則合成装置

##### (57) 【要約】

【課題】 本発明は大規模な計算機を必要とすることなく、音声素片を生成する際に必要なラベリング処理を行なうとともに、1つの計算機システム上に、音声信号を処理して音声情報を作成する回路と、音声合成を行なう回路とを構築して、音声素片の生成を容易にし、合成音声の話者変更を容易にする。

【解決手段】 曲再生回路2によって予め登録されている曲の歌詞を表示して、メロディ音を出し、話者に歌を歌わせながら、音声情報抽出回路3によって前記話者の歌音声を取り込み、この歌音声と前記曲が持つ各音符の長さ、テンポなどの情報に基づき、前記歌音声を音節毎に分割し、各音節毎にラベリングを行なって音声素片を作成し、音声の合成対象となるテキストデータが入力されたとき、このテキストデータに応じて音声素片を組み合わせて合成音声を生産する。



【0003】上記各構成部について詳細に説明すると、音声情報抽出回路102は、話者が発生した音声を取り込み、この音声信号を分析して種々の音響パラメータ（声の大きさ、スペクトルなど）を抽出する分析部10

5と、この分析部105から出力される各音響パラメータを取り込みこれら各音響パラメータの変化量などを計算する変化量計算部106と、この変化量計算部106から出力される各計算結果（変化量）および入力された音韻情報に基づき音響パラメータの変化量が大きい箇所を検出してこの部分を音声の境界（音節）と判定して音波形を区切るとともに、これらの各単位波形に対して自動的にラベリングを行うセグメンテーション部107とを備えている。

【0004】合成音声の元となる話者の音声データベースを作成するときは、話者が発声した音声を取り込み、この音声を分析して各種の音響パラメータを抽出した後でこれら各音響パラメータの変化量を計算してこの計算結果及び発声した音声に対応する音韻に基づき前記音声の境界を判定して自動的にラベリングを行い、これによって得られたラベリング済みの音声データを前記音声データベース103に登録する。音声データベース103は、大きな記憶容量を持つハードディスク装置、光磁気ディスク装置等によって構成されており、音声情報抽出回路102からラベリング済みの音声データが出力される毎にこれを取り込んで累積記憶し、音声合成回路104から読み出し指示がなされる毎に、指定された音声データの音声合成回路104に供給する。音声合成回路104は、前記音声データベース103に登録されているラベリングされた音声データから音声素片を作成する音声素片生成部108と、音声の合成対象となるテキストデータが入力されたときに該テキストデータに応じて音声素片生成部108で作成された各音声素片を組み合わせさせて合成音声を生成する音声合成部109とを備えている。音声合成動作に先立って、音声素片生成部108は音声データベース103に累積記憶されているラベリングされた音声データから合成音声に必要な音声データを取り込み、この音声データから音声素片を作成しておく。前記合成音声に必要な音声データは音声素片の作成方法によって種々異なるので、必要とするラベリングされた音声データの選択は音声の研究者や専門家の手作業により行われる。そしてテキストデータが入力されたときこのテキストデータに応じて前記各音声素片を組み合わせさせて合成音声を生成する。

40 【０００５】

【発明が解決しようとする課題】しかしながら、上述した従来の音声規則合成装置１０１においては、ラベリング処理を行なう際、入力音声に対して膨大な演算処理を行わなければならないことから、大規模な計算機を必要とし、これによってシステム全体のコストが高くなるという問題があった。また、このような音声規則合成装置１０１では、音声情報抽出回路１０２と、音声合成回路１０４とを別々のシステムによって構成しているの

50 り直しや音声素片の再作成など、煩雑な操作を再度行な

わなければならないとともに、これらの操作に専門の知識を必要とし、さらに手作業で音声素片の作成を行わなければならない。このため、合成音声の元になる話者の変更を行なうのが難しいという問題があった。本発明は上記の事情に鑑み、音声信号を解析するための大規模な計算機を必要とすることなく、音声素片を生成する際に必要なラベリング処理を行なうことができるとともに、1つの計算機システム上に音声信号を処理して音声情報を作成する音声情報抽出回路と、音声合成を行なう音声合成回路とを構築して音声素片の生成を容易にし、合成音声の話者変更を容易にすることができる音声規則合成装置を提供することを目的としている。

#### 【0006】

【課題を解決するための手段】上記の目的を達成するために本発明による音声規則合成装置は、請求項1では、予め登録されている曲を再生する曲再生回路と、この曲再生回路で再生された曲に合わせて入力された話者の歌音声を取り込むとともに、前記曲再生回路で再生された曲の楽譜に基づき、前記歌音声を処理して音声素片を作成する音声情報抽出回路と、入力されたテキストデータに基づき、音声情報抽出回路で作成された音声素片を組み合わせて音声を合成する音声合成回路とを備えたことを特徴としている。また、請求項2では、請求項1に記載された音声規則合成装置において、音声音声情報抽出回路は、前記曲再生回路で再生された曲のメロディに合わせて歌う、前記話者の歌音声を前記メロディの楽譜に合わせて音節単位で区切って音声素片を作成することを特徴としている。

#### 【0007】

【発明の実施の形態】以下、本発明を図面に示した形態例に基づいて詳細に説明する。図1は本発明による音声規則合成装置の一形態例を示すブロック図である。この図に示す音声規則合成装置1は、曲再生回路2と、音声情報抽出回路3と、音声合成回路4とを備えており、曲再生回路2によって予め登録されている曲の歌詞を表示してメロディ音を出し、話者に歌を歌わせながら、音声情報抽出回路3によって前記話者の歌音声を取り込み、この歌音声と前記曲が持つ各音符の長さ、テンポなどの情報に基づき、音節毎にラベリングを行なって音声素片を作成し、音声の合成対象となるテキストデータが入力されたとき、このテキストデータに応じて音声素片を組み合わせて合成音声を生成する。

【0008】曲再生回路2は、図2(a)に示す如くいろは歌など、50音を効率良く入力させるのに必要な歌詞データおよび歌詞がしっかりと歌えるようにゆっくりしたメロディを持つ楽譜データが格納されるハードディスク装置5と、このハードディスク装置5に格納されている歌詞データおよび楽譜データを読み出し、この歌詞データに基づいた表示データを生成するとともに、前記楽譜データに基づいたメロディ音信号を生成する曲再生

部6と、この曲再生部6から出力されるメロディ音信号を取り込み、メロディ音を発声するスピーカ7と、前記曲再生部6から出力される表示データを取込み、歌詞を表示するディスプレイ8と、前記曲再生部6から出力される楽譜データを取り込んで、音符の長さに対応したセグメンテーションデータ（楽譜を構成する各音符の継続時間を示すデータ）を生成する楽譜データ変換部9とを備えている。話者の歌音声を取り込むとき、予め登録されている歌詞データと、この歌詞データに対応する楽譜データとに基づいて歌詞を画面表示して、前記話者に見せながらメロディ音を発生してこれを前記話者に聞かせるとともに、前記楽譜データに基づきセグメンテーションデータを生成してこれを音声情報抽出回路3に供給する。

【0009】音声情報抽出回路3は、図2(b)に示す如く前記話者が発した歌音声を取り込み歌音声信号を生成するマイクロホン10と、このマイクロホン10から出力される前記歌音声信号の記録エリアとして使用されるハードディスク装置11と、前記マイクロホン10から出力される歌音声信号を取り込んでデジタル化しながら前記ハードディスク装置11に記憶させた後、読み出し指示があったとき前記ハードディスク装置11に記憶されているデジタル化された歌音声信号を読み出す録音部12と、前記曲再生回路2の楽譜データ変換部9から出力されるセグメンテーションデータで示された曲のテンポおよび音符の長さに基づき曲の時間経過に応じて録音部12によって読み出された歌音声信号を音節単位に分割してラベリングを行なう音声分割部13とを備えている。

【0010】前記曲再生回路2によって歌詞が表示されるとともにメロディ音が出され、これに対応して話者が歌を歌っているとき、この歌音声を取込んで歌音声信号に変換しながら前記曲再生回路2から出力されるセグメンテーションデータで示された曲のテンポおよび音符の長さに基づき、曲の時間経過に応じて前記歌音声信号を音節単位に分割するとともにラベリングを行なって音声素片を作成し、これを音声合成回路4に供給する。

【0011】音声合成回路4は、音声情報抽出回路3の音声分割部13から出力される各音声素片を取り込んで蓄積する音声素片部14と、音声の合成対象となるテキストデータが入力されたときこのテキストデータに応じて音声素片を組み合わせて合成音声を生成する音声合成部15とを備えており、音声分割部13から出力される各音声素片を取り込んで蓄積し、音声の合成対象となるテキストデータが入力されたときこのテキストデータに応じて音声素片を組み合わせて合成音声を生成する。

【0012】このように、この形態例では曲再生回路2によって予め登録されている曲の歌詞を表示してメロディ音を出し、話者に歌を歌わせながら音声情報抽出回路3によって前記話者の歌音声を取り込み、この歌音声と

前記曲が持つ各音符の長さ、テンポなどの情報に基づいて歌音声を音節毎に分割し、各音節毎にラベリングを行なって音声素片を作成し、音声の合成対象となるテキストデータが入力されたときにこのテキストデータに応じて音声素片を組み合わせて合成音声を生産するようにしたので、音韻情報などを入力することなく、入力された音声を音節単位に分割して、ラベリングを行ない、音声素片を作成することができる。

【0013】これによって、音声信号を解析するための大規模な計算機を必要とすることなく、音声素片を生産する際に必要なラベリング処理を行なうことができるとともに、1つの計算機システム上に音声信号を処理して音声情報を作成する音声情報抽出回路3と、音声合成を行なう音声合成回路4とを構築して、音声素片の生成を容易にし、合成音声の話者変更を容易にすることができる。また、上述した形態例においては、ハードディスク装置5内に、いろは歌など、50音を効率良く入力させるのに必要な歌詞データおよび歌詞がしっかりと歌えるようにゆっくりしたメロディを持つ楽譜データを格納するようにしているが、他の大容量記憶装置、例えばフロッピーディスク装置、光磁気ディスク装置などに、歌詞データや楽譜データなどを格納するようにしても良い。この際、50音が出現する曲であれば、上述したいろは歌以外の曲の歌詞データや楽譜データを格納しても良い。

#### 【0014】

【発明の効果】以上説明したように本発明によれば、音声信号を解析するための大規模な計算機を必要とすることなく、音声素片を生産する際に必要なラベリング処理を行なうことができるとともに、1つの計算機システム

\* 上に、音声信号を処理して音声情報を作成する音声情報抽出回路と、音声合成を行なう音声合成回路とを構築して音声素片の生成を容易にし、合成音声の話者変更を容易にすることができる。

#### 【図面の簡単な説明】

【図1】本発明による音声規則合成装置の一形態例を示すブロック図である。

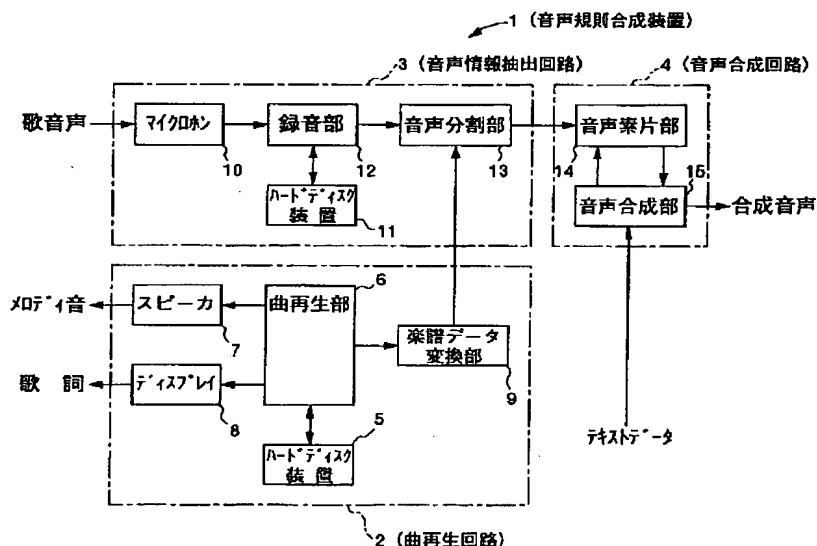
【図2】(a)及び(b)は図1に示す曲再生回路のハードディスク装置内に格納される歌詞データ、楽譜データの一例および入力される歌音声の波形例を示す模式図である。

【図3】従来から知られている音声規則合成装置の一例を示すブロック図である。

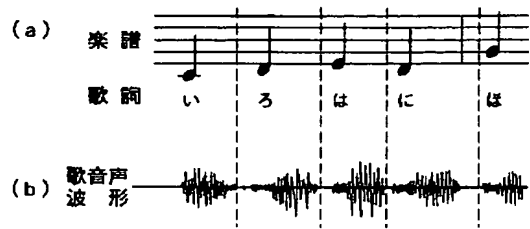
#### 【符号の説明】

- 1 音声規則合成装置
- 2 曲再生回路
- 3 音声情報抽出回路
- 4 音声合成回路
- 5 ハードディスク装置
- 6 曲再生部
- 7 スピーカ
- 8 ディスプレイ
- 9 楽譜データ変換部
- 10 マイクロホン
- 11 ハードディスク装置
- 12 録音部
- 13 音声分割部
- 14 音声素片部
- 15 音声合成部

【図1】



【図2】



【図3】

